Predicting wine quality based on physicochemical attributes

Zhihao Xie, Natalie Trusdell, Oswin Gan & Sam Lokanc

2025-12-06

Table of contents

Introduction

Quality assessment is an essential step in the wine making process as it helps ensure consistency and consumer satisfaction. The quality of wine has historically been evaluated by wine makers and sommeliers, who leverage their training and expertise to determine if a product meets acceptable standards Langstaff (2010). While there have been advancements in the detection of the physicochemical properties of wine and how they relate to flavour Polášková, Herszage, and Ebeler (2008), quality assessment remains a subjective metric.

Here we ask if wine quality can be predicted using machine learning. Our focus is to create an interpretable linear regression model trained on physicochemical wine data to perform this task. The utility of such a model is two-fold as it can be used as a tool to help less experienced wine makers and sommeliers get a sense of the quality of a given wine. Furthermore, the model would help to provide a more objective framework for wine assessment through interpretation of its coefficients. To achieve this, we will train our model on data related to vinho verde white

wines from the Minho region of Portugal which contains the physicochemical and sensory data of 4898 wines Cortez et al. (2009).

Methods

Data

Analysis

Results and Discussion

References

Cortez, Paulo, António Cerdeira, Fernando Almeida, Telmo Matos, and José Reis. 2009. "Modeling Wine Preferences by Data Mining from Physicochemical Properties." *Decision Support Systems* 47 (4): 547–53.

Langstaff, SA. 2010. "Sensory Quality Control in the Wine Industry." In *Sensory Analysis for Food and Beverage Quality Control*, 236–61. Elsevier.

Polášková, Pavla, Julian Herszage, and Susan E Ebeler. 2008. "Wine Flavor: Chemistry in a Glass." *Chemical Society Reviews* 37 (11): 2478–89.